Guidelines for Industrially-Based Multiple Case Studies in Software Engineering

J.M. Verner
CSE, UNSW
Sydney, Australia
june.verner@gmail.com

J. Sampson
Epsis AS
Bergen, Norway
jsa@epsis.no

V. Tosic
NICTA & UNSW
Sydney, Australia
Vladimir.tosic@ nic-ta.com.au

N.A. Abu Bakar
CSE, UNSW
Sydney Australia
nurazzah@uum.edu.my

B.A. Kitchenham
Keele University
Staffordshire, UK
B.A.Kitchenham @cs.keele.ac.uk

***Abstract*—Without careful methodological guidance, case studies in software engineering are difficult to plan, design and execute. While there are a number of broad guidelines for case study research, there are none that specifically address the needs of a software engineer undertaking multiple case studies in an industrial setting. Through a synthesis of existing best practices in case study research, we provide a set of comprehensive guidelines for conducting multiple case studies in software engineering research. Our guidelines can assist software engineering researchers with all stages of multiple case study research, although in this paper we concentrate on the early phases, such as focusing the case study and detailed plan design. To date, three exploratory research projects found our guidelines very useful. We illustrate our guidelines with examples from one of these projects.**

Keywords-case study guidelines; software engineering; case study planning; case study design; case study data collection

I. INTRODUCTION

Rather than using a large number of samples and following a rigid protocol to examine a limited number of variables, case study methods involve an in-depth examination of a single case or a small number of cases (comparative case study). They provide a systematic way of looking at events, collecting data, analysing information, and reporting the results. As a result, case study researchers may gain a greater understanding of why something happened as it did, and what might be important to investigate in future research. Case studies can be used for both generating and testing hypotheses [13]. In software engineering, case studies are recognized as a formal research method, and are becoming popular not only for evaluating research, but also to observe, explain and explore phenomena in real-life settings.

Yin suggests that the case study should be defined as a research strategy–an empirical inquiry that investigates a phenomenon within its real-life context [45]. While all case studies include qualitative evidence, many case studies also present quantitative evidence [6, 8, 18, 20, 21, 23, 24, 27, 29-32, 35, 39, 40]. Case studies can be very diverse. For example, there are significant differences between single case studies (with one case) and multiple case studies (with several cases) [25]. Case study research may be classified based on its purpose into descriptive, explanatory, exploratory or evaluatory [33, 37, 40]. Irrespective of these and other differences, all case studies rely on multiple sources of evidence and benefit from the prior development of theoretical propositions [44, 45]. In the case of exploratory research it is important to use industrially-based cases as the context from which a theory or artifact may emerge [10, 40]. Kitchenham et al. [19, 20, 21] discuss case studies for industrial evaluation of software technology.

Although case studies are a very useful research method for software engineering, software engineering case study methodology is relatively immature, with case studies often being poorly performed. The method is difficult to apply rigorously [4] and, despite its popularity, many problems are evident in papers reporting results. In particular, in a thorough review of prominent case studies in software engineering, computer science and information systems, we determined issues related to poor planning, design and reporting of case study research. While there are many broad guidelines and checklists for case study research (e.g., [1, 7, 9, 15, 19, 42-45]), none wholly address, in detail, the needs of software engineers undertaking multiple case studies in industry settings. Many of the guidelines are either incomplete, somewhat vague and/or do not provide useful concrete examples [15, 44, 45]. Hence, the motivation for our research is to provide a detailed step-by-step guide specifically for researchers in software engineering.

Our detailed set of guidelines can help with the process of planning, constructing and reporting on all types of multiple case studies. Our focus is on multiple, industry-based studies that describe, evaluate, explore or explain a contemporary software engineering phenomenon within its real-life context. Our guidelines are also appropriate for single case study design. The guidelines presented here were developed mainly for the use of Ph.D. students involved with exploratory research in software engineering.

We organized our case study guidelines into several phases generally seen in case study research (but not necessarily documented in the case study literature). The phases are: research initiation or pre-planning, administration, focus case study or planning, design case study plan, data collection, data analysis (including evaluation and conclusions), and reporting.

9781-4244-2865-6/09/$25.00 ©2009 IEEE

We also recommend updating research notes and database as a continuous activity. These phases are organized as shown in Figure 1, which is a process model in the Business Process Management Notation (BPMN) 1.0. BPMN [38] is a relatively simple flow-chart based notation for graphical description of sequences of business activities and supporting information (e.g., documents).

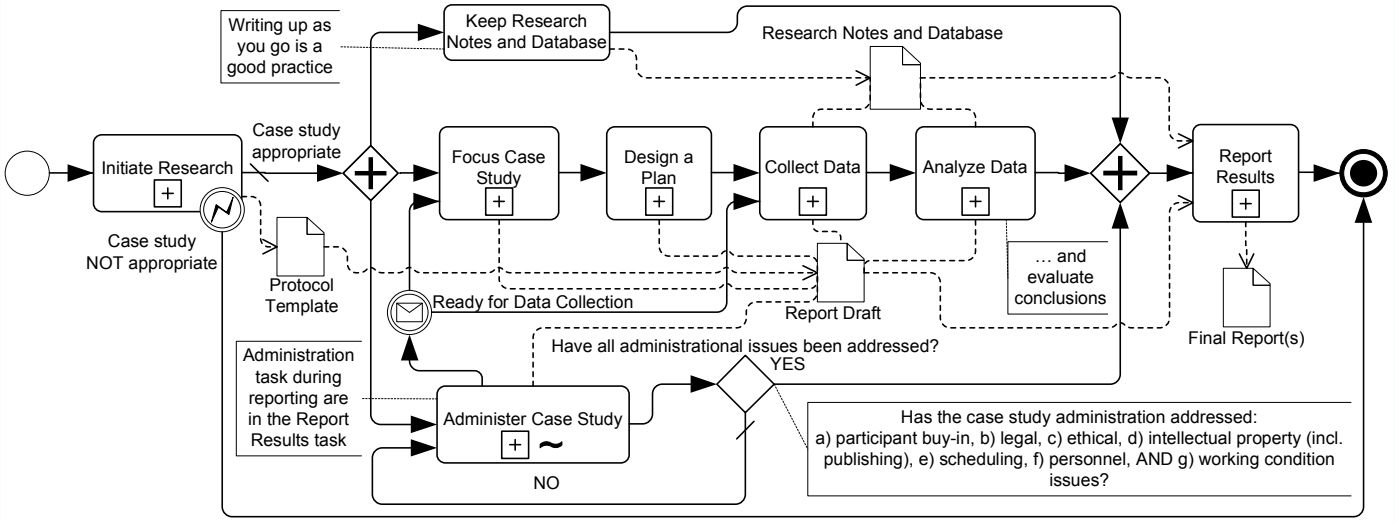

Figure 1. General organization of phases in case study research (BPMN 1.0 diagram)

There are four main categories of elements in BPMN: flow objects, connectors, artifacts, and swimlanes. The standard types of flow objects are: event (represented as a circle), activity (rounded rectangle), and gateway (diamond). All of them can contain optional symbols (markers) that differentiate between subtypes (e.g., a sub-process is represented as an activity with the boxed '+' sign in the bottom middle, a 'parallel fork/join' gateway is represented with a '+' sign within a diamond). In Figure 1, our case study phases are shown as sub-processes. There are also comments (open rectangles).

In this paper, we concentrate on explaining in detail the early phases of case study research (which are often not given due attention by researchers). We identify steps within each phase, but note that ordering of steps within a phase need not always be sequential. We provide some concrete examples (from actual case study research in which we have been involved) to assist in a better understanding of our guidelines. Due to space constraints data analysis and reporting are described only at a fairly high level. In Section II we discuss research initiation or pre-planning, in Section III the case study focus, in Section IV design of the plan, in Section VI data collection, in Section VII data analysis, and in Section VIII reporting. Section IX discusses further work.

II. Research Initiation or Pre-planning

The first phase ensures that at the outset of a project sufficient preparation and training is undertaken by the case study investigators. Its four steps are: defining research objectives, performing a comprehensive literature review, deciding if a case study is appropriate, and setting up a preliminary case study protocol template.

1. Define the study objectives

In this preparatory step, broad and overarching objectives for the research project are defined. In general, they will be a set of high-level objectives for the study based on the identified area of interest [11].

Example: *"The objective of this study is to investigate factors leading to successful medical expert systems, given that most medical experts systems fail shortly after their introduction into the workplace. This study is classified as an exploratory case study as it aims to explore a subject area which is not well understood.*"

Step 2. Undertake comprehensive literature review

A comprehensive literature review and analysis with sufficient breadth and depth is used to form a solid foundation for the research. Only once an extensive prior art study has been completed can a researcher successfully identify where additional contributions are possible. The literature review should consider all previous significant work in the area, i.e., identify the key authors and their contributions, any significant position papers, and previous case studies. A key output from this step should be a set of statements describing the necessity for further investigation in the area of study. Documenting the motivation for further investigation allows the researcher to narrow the scope of the planned research, and helps to ensure that the study will be possible given time and resource constraints.

Step 3. Decide if case study is appropriate

The researcher must determine if a case study is appropriate given the findings from the literature and research objectives. In general, using the case study research method is ap-

propriate when “how” or “why” research questions are proposed, when the researcher has limited control over the events, and when the focus of the research is to study a phenomenon within some real-life context [45]. Kitchenham et al. [19] suggest that we need case studies in software engineering “to evaluate not only how or why, but also which is better”. If a case study is deemed appropriate the researcher should provide an explanation of the intended case study research and justification for choosing this approach. The justification should address the strengths and weaknesses of the case study approach (e.g., [5, 45]).

Step 4. Set up the protocol template

One key output of the preplanning phase is a research protocol template [4, 12]. The protocol provides a written record of the procedures followed throughout the case study. It should include aspects of data collection, storage and the rationale for the research. It should also provide a description of case selection, present the case study framework and discusses methodological considerations. A code of conduct should be described in the protocol along with procedures that address areas such as introductory letters, contact person/case manager, and administration issues such as public availability of interview material.

The protocol is a living document that cannot be completed in the first phase of case study research but should be updated as the research progresses. At this stage, only the protocol template can be determined and the first few sections completed in draft form. The research objectives and the literature analysis are added to the protocol once completed. During subsequent phases, the objectives are refined and the literature analysis updated. At the completion of every phase of a case study, the case study protocol template should be updated. The case study protocol will be complete only when all phases are completed. The protocol should be included as part of the research report as an appendix. Examples of case study protocol templates can be found at [3, 4].

The protocol template can increase reliability of case study research [45]. In particular, Yin [45] remarks that having a case study protocol is essential when undertaking multiple-case studies. A desirable training task is that all case study investigators participate in the development of the protocol [44, 45]. Any modifications to the protocol should be discussed with those involved before the modifications are made. In addition, a protocol can ensure that any researchers wishing to replicate the study can do so in the knowledge that they are following exactly the same protocol as that used in the original research. Brereton et al. [4] discuss the development of protocol templates for planning systematic reviews.

III. Case Study Administration

Overall, administration plays an important role in ensuring that issues related to a) legal, b) ethical, c) intellectual property (including publishing), d) scheduling, and e) working conditions, are met. There are five main activities or steps in this phase, but administration is an ‘ad hoc’ sub-process, in the sense that its activities can be done in any order, with any frequency, and in parallel with almost any other phase.

1. *Review the need for legal agreements*

Where necessary, legal agreements and memoranda of understanding should be created. It is essential that legal agreements are formulated at the outset of the project.

2. *Organise requirements for ethics*

Requirements for ethics must be established and ethical clearances formalized before data collection starts. These requirements are documented in the case study protocol. Informed consent must be obtained from organizations and individuals before data is collected [15].

3. *Find out what the publishing criteria is*

Establishing publishing criteria facilitates that all involved parties understand the intent to publish and agree where, when, to what extent the material will be published, and what review procedures will be required before submission.

4. *Ensure familiarity with partner schedules*

In dealing with industrial partners and their schedules, the researchers must ensure that any schedule issues are formalized and documented in the protocol.

5. *Oversee working conditions*

Any required working conditions and case study administration details are documented in the protocol.

IV. Case Study Focus or Planning

In this section we describe in detail the 16 steps that comprise the phase that determines basic characteristics of the case study. This phase is sometimes called planning but we find the term “case study focus” more appropriate, because a detailed plan will be developed fully in the subsequent phase. The order of the 16 steps need not be sequential in all circumstances.

Step 1. Refine research objectives and formalize research questions

The focus of the study can be established by refining research objectives into questions about the situation or problem to be studied and the purpose of the study. The research questions must be precise and unambiguous, not have been addressed by previous literature, achievable with the case study approach and not better suited for another scientific approach.

Case study research generally answers one or more questions which begin with "how" or "why." The questions are targeted at considering a limited number of events or conditions and their inter-relationships. To assist in targeting and formulating the questions, the literature review conducted in the research initiation phase established characteristics of past research, which leads to refined, insightful problem questions. A careful definition of questions pinpoints where to look for evidence and helps determine the analysis methods to be used in the study. We recommend the researcher to explicitly state the research questions in this step.

Example: *“The overarching research question is: Why do most medical expert systems fail shortly after their introduction, while some clinical laboratory expert systems last for longer than five years and continue to form an important part*

of a user's job?"

Within the overarching research question several sub-questions need to be addressed. For the above example, some sub-questions are:

RQ1. How do motivational factors encourage longevity in clinical expert systems?

RQ2. How do long-lasting clinical laboratory expert systems support maintenance in order to adapt to changes over time?

Step 2. Define clear research proposition(s)

Propositions are predictions about the world that may be deduced logically from theory [34]. Dubin [10] remarks that the usual form of propositions is the "IF ... THEN ..." format, but not all researchers agree on this (this format is not always the clearest; furthermore it implies causal relationships, while in some cases only correlation of facts can be determined). According to Yin [44], a proposition "directs attention to something that should be examined within the scope of the study". The researcher must ensure that every research proposition is: a) precise and unambiguous; b) achievable with the case study approach; and c) corresponds to the defined research objective(s). Research propositions should be placed together with their corresponding research question to make it clear how they have been derived.

For the example from the previous step, we have:

RQ1. How do motivational factors encourage longevity in clinical expert systems?

PR1.1: IF there are motivational factors THEN users will continue to use the clinical expert system.

RQ2 How do long-lasting clinical laboratory expert systems support maintenance in order to adapt to changes over time?

RP2.1: Long-lasting clinical expert systems are easy to maintain.

RP2.2: Domain experts are involved in maintenance activities of long-lasting clinical expert systems.

RP2.3: In long-lasting clinical expert systems, knowledge is easy to acquire from domain experts.

RP2.4: In long-lasting clinical expert systems, knowledge is easy to represent.

RP2.5: Long-lasting clinical expert systems show significant increase in the size of their knowledge base since their first deployment.

Step 3: Identify unit of analysis

At this stage the researcher needs to decide what a case will consist of. That is, the researcher should decide whether a case is a project, a department, a set of different projects, an organization, etc. Will the study focus on individuals, groups or an entire organization? Is the unit of analysis a project, a piece of software, or a decision? To decide, the researcher must closely examine the research questions. These often indicate an appropriate unit of analysis. Finally, the researcher should consider what generalizations are hoped for at project completion. Does the researcher hope to generalize to other projects, organizations, individuals or decisions? The researcher needs to consider exactly how the research questions can be answered. This will involve determining what forms are necessary, planning and preparing the interview protocol, what types of equipment will be needed, (e.g., recording of data), starting to plan the analysis, what forms and structure will be used, and will software be required to help with this.

Step 4: Define conceptual framework

In Step 2 of the research initiation phase, a comprehensive literature review was performed by the researcher. From the literature review the researcher should gain an insight into prior research and have the means to construct a list of key points and factors that can be used as a basis for data collection. These key points become the foundation of a conceptual framework that will evolve over time as the research is carried out and the findings analysed. For example, a preliminary conceptual framework for our earlier example, consisting of factors extracted from the expert systems literature, is shown in Figure 2.

Step 5: Define concepts and measures

A concept is an idea being investigated; e.g., from our example, "ease in maintenance", "growth in size of the knowledge base", are concepts that should be defined, whereas a measure actually quantifies "ease in maintenance" or "growth in size of the knowledge base". Some concepts such as "ease" or "easy" obviously need to be refined to much greater detail in order to be measured. At this stage, the researcher should ensure that every defined measure is: a) precise and unambiguous; b) relevant for the concept(s) it is supposed to measure; c) measurable within the context of the case study; d) with a precisely determined measurement unit (in the case of several possible measurement units, it is advisable to determine one of them to be used throughout the case study); e) illustrated with examples for every concept (this improves understandability). The concepts found in the research propositions should be defined and added to a research glossary; whereas the measures should be defined and made explicit in a table included in the main body of the research methodology section. The researcher should make explicit the concepts and measures to be analysed or evaluated. It may happen that although a concept is interesting and should be defined – it may not be appropriate to measure and analyse it explicitly. For example, the concept "knowledge" in an expert system may be an important concept in a case study about expert system success. However, it may not be measured directly; rather it is the participants' interpretation of the "knowledge" that can only be described. With this in mind, the researcher should systematically go through each defined concept and measure and state explicitly whether or not it will be analysed.

Step 6: Define how you will analyse results

Furthermore, it is important that the researcher has a plan for analysing the defined concepts and measures. By looking back at the concepts and measures, the researcher should now decide exactly how they will be analysed. A tabular format may be useful to represent each concept and measure, and the planned analysis for each. At this stage, it is sufficient to state, at a relatively high level, how the concepts and measures will be analysed, but it is important that the researcher plans for

specific and sufficient data for an analysis. For example, a number of the "concepts" identified, defined and selected in step 5, may be analysed using qualitative data analysis techniques, such as thematic coding and rich interpretation; whereas the researcher may plan to analyse the "measures" using simple quantitative techniques, Likert scales, or more sophisticated statistical methods.

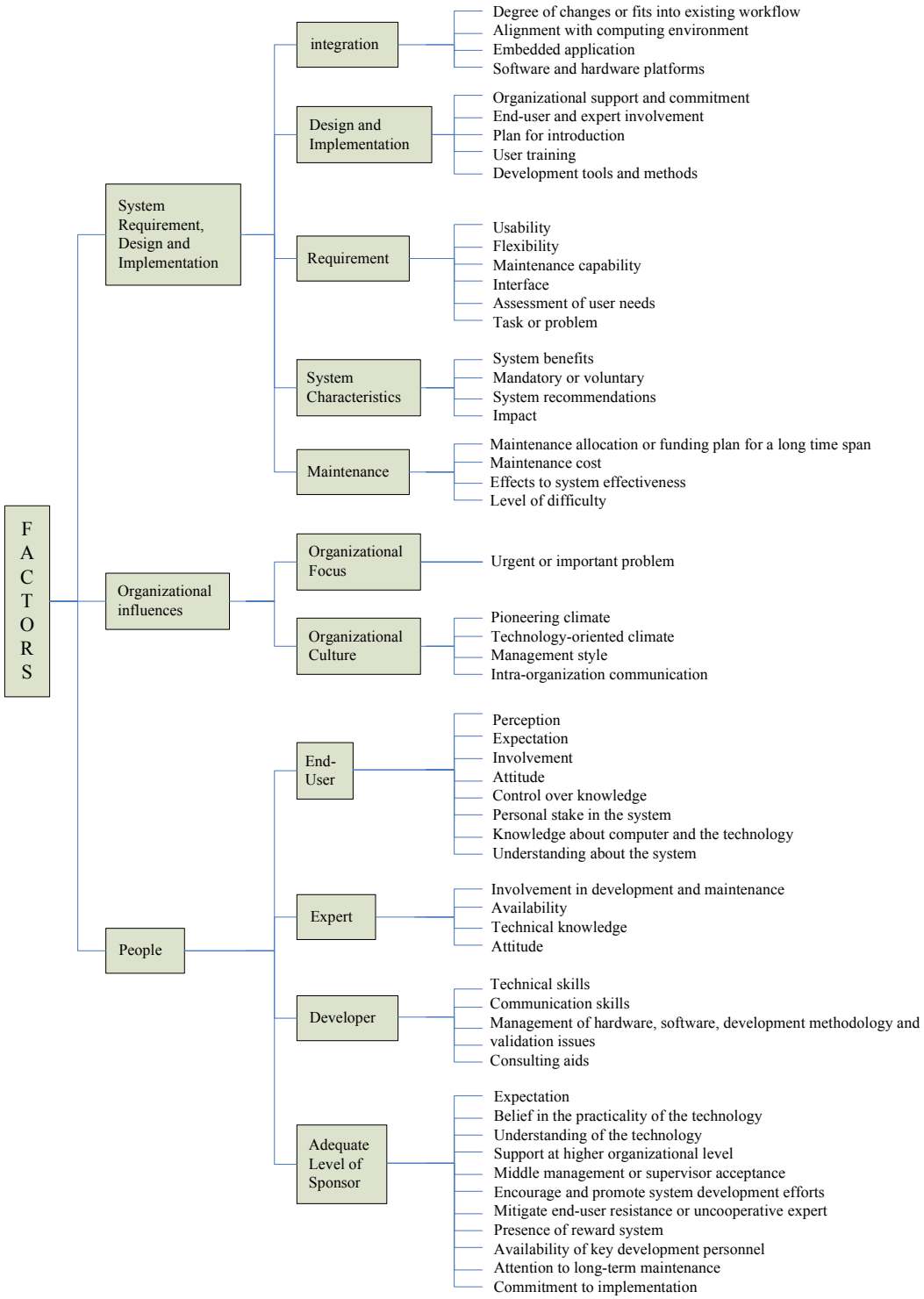

Figure 2: Proposed framework developed from the literature

For each planned analysis the researcher may provide a description of how a particular analysis will be undertaken, e.g., using a comparative analysis of case study participants responses to a particular question or set of questions.

Step 7: Identify characteristics of cases to study

A general description of a case should have been defined in step 3, but now further details should be provided. For example, if the case is a department, its characteristics may include "*a department in the healthcare environment using a clinical laboratory expert system*".

Step 8: Identify & justify sites and/or individuals that are the focus

Yin [44], proposes two criteria for selecting sites: first, sites where similar results are predicted (this may be useful for "literal" replications); second, sites may be chosen for theoretical replications. That is chosen such that contradictory results are predicted. With careful site selection the researcher can extend and revise the initial propositions of the study. Site selection should be carefully thought out rather than be opportunistic. This is the ideal, but it may not be possible. Researchers often have to take what they can get and consider themselves lucky if they can get enough cases that fit their criteria. A researcher may begin site selection by considering the nature of the topic. Research on organizational level phenomena would require site selection to be based on the characteristics of the organizations. These may include industry, company size, organizational structure, profit or not for profit status, public versus private ownership, geographical coverage, degree of vertical or horizontal integration, how the researcher will get access, etc. Researchers in specific technologies, methodologies or organizational structures should consider these characteristics when selecting sites. Once the limiting factors are determined specific sites may be identified and approached.

Step 9: Define boundary of case study

Describe the scope of the study, and what criteria are used to determine its scope [36].

For example "*opinions will be gathered only from individuals with first-hand experience with clinical laboratory expert systems, such as knowledge engineers who developed them, domain experts who contributed to the development and users of the system"*.

Step 10: Identify baseline for evaluation if appropriate

This step is optional and is only required if the researcher intends to conduct a comparative study. For example if a researcher wishes to compare the effect of a new programming language they need to know what the standards were before the new language was introduced (but see step 2 of the next, design, phase for some caveats).

Step 11: Establish benefits for organization or group (includes buy-in for appropriate groups)

Approaching a potential site is a crucial point in orchestrating a case study research project. Here again the topic of the study is a key to determine whom to contact. The researcher must identify an individual with enough authority to approve the project. She/he should prepare carefully, particularly if having to place a cold call or writing to the organization. She/he should clearly describe the project and who will be involved – researchers, assistants and company employees. The contact should be told the amount of time, effort, and expense required from the organization, and what is in it for the organization should they participate in the research. An organization should not be harmed by its participation, and the organization and the employees must know that the researcher will not betray their confidence. On the other hand the researcher must seek assurance that reasonable openness will be provided and that essential data will be made available.

Benefits to a participating organization are varied. They may include learning more about the organization, or getting feedback on new insights from the researcher. There is also the opportunity to contribute to knowledge and research. The organization may, or may not, want to be identified when the research is published, and if it does wish to be identified there is the additional benefit of recognition and publicity.

Step 12: Identify feasible cases

Does every proposed case: a) have sites and/or individuals that are the focus of the case study; b) lie within the defined boundary of the case study; c) have a baseline for evaluation (if appropriate); d) satisfy all other relevant constraints. If any criterion is not met, the researcher must return to the corresponding step in this phase to perform further revision (see [36]). Are there any constraints that need to be taken into account? For example, will the organization insist that the researcher is accompanied by a senior manager in all interviews and will the fact that a third party is present affect the results?

Step 13: Select cases to study

Every selected case for the study should satisfy that: a) it is precise and unambiguous; b) is it likely that the required information will be available given the environment and organization in which the case is embedded; c) it corresponds to the defined/refined research objective(s) and research proposition(s); d) it is observable and measurable as required during the time frame of the planned research.

Step 14: Select pilot case(s)

It is highly recommended that the researcher carries out a pilot study before undertaking the "real" cases. According to Yin [42-45], a pilot case study helps the researcher prepare for data collection and refine data collection plans. The researcher may trial a broader data collection plan during the pilot study. The selection criteria for a pilot case study are usually convenience, access and geographic proximity. The researcher should aim to select a pilot site where a congenial relationship between interviewees can be established and where the participants are aware that the researcher is still at an early stage of research. When the research is highly exploratory a single case may be useful as a pilot study. The goal will be the unit and will familiarize the researcher with the phenomenon in its context [2].

Step 15: Ensure case study researchers have appropriate skills

and understand their purpose and role

Does every researcher in the case study: a) understand her/his purpose and role; and b) has appropriate skills? If any criterion is not met, appropriate education/guidance should be provided to the researcher. If this is not possible, the researcher should be replaced, or the researcher should return to the (re)definition of his/her research objectives.

Step 16: Decide if you will get appropriate level of confidence

The case study researcher should determine if the number and type of cases selected will provide an appropriate level of confidence. However, a case study researcher is usually limited with respect to locating willing organizations, resources, and expense. This usually means that only a limited number of cases are indeed available and are possible with respect to time. It may be argued that "strong" conclusions based on such a limited sample of the population, are not valid. In survey research this is known as "sampling error." However, in many cases it is sufficient for case study researchers to report a set of findings applicable only to the chosen selection of cases. The case study researcher should aim not to generalize over a wider population but instead may represent the findings using rich interpretation methods and through a conceptual framework, which may then be validated by further empirical research.

V. Design a Detailed Plan

Yin [45] remarks that "the design is the logical sequence that connects the empirical data to a study's initial research questions and, ultimately to its conclusions". The following nine phases comprise case study design in our guidelines.

Step 1. Convert propositions to hypotheses

A hypothesis is an empirically testable statement that is generated from propositions defined in the planning phase. See example in Table 1. According to Shanks and Parr [34], values of concepts that appear in propositions should be defined before empirical testing is conducted. In addition, it is possible that one proposition may generate many hypotheses. A very useful outcome of hypotheses generation is that the researcher refines the definitions of the concepts and measures more precisely so they can be evaluated.

Table 1. Example proposition and hypothesis

Proposition	**Hypothesis**
PR1.1 (for RQ1): IF there are motivational factors THEN users will continue to use the clinical laboratory expert system.	**H1.1**: Motivating factors (such as: system is easy to use, system fits into existing workflow, system fits into computing environment, tasks are appropriate for the user's work satisfaction), are highly correlated with the continued use of a clinical laboratory expert system.

First, the researcher must identify empirical indicators for the terms in the propositions. This is called operationizing a concept, as it makes the concept ready to be used. For example, if the researcher considers positive impact, then there is a need to define what this is and how it will be measured.

The hypothesis must be expressed in a form suitable for refutation. In the previously discussed example of clinical expert systems, we had research proposition PR1.1 (see Table 1) and the subsequent steps of the previous phase clarified that examples of motivational factors are *system is easy to use, system fits into existing workflow, system fits into computing environment, tasks are appropriate for the expert's work satisfaction.* If we have empirical indicators for "*easy to use*", "*system fits into existing workflow", "system fits into computing environment" ,"tasks are appropriate for the expert's work satisfaction"* and "*continued use*", we can then decide whether or not to accept the hypothesis for this example.

Step 2. Identify method of comparison if appropriate

If the case study is an evaluation of a particular method or a particular tool, then a method for contrasting the results of using one method (or tool) with the results of using another must be identified. This is essential to help avoid bias and ensure internal validity. Kitchenham et al. [19], note that there are three ways to facilitate this comparison when undertaking a method evaluation in a case study setting:

- Select a sister project with which to compare.
- Compare the results of using the new method against a company baseline.
- If the method applies to individual components, it should be applied at random to some product components and not to others.

Validation exercises based on implementing an existing system using a new methodology can provide useful information about the advantages and limitations of new technologies. However, they cannot provide information about potential productivity or timescale advantages since examples are seldom produced under the strict process controls required for commercial products [4].

Step 3. Minimize the effect of confounding factors

Confounding factors are risk factors that may affect the results of a study. They play an important role in the design of case studies when a researcher is studying a phenomenon in a real-world setting. Confounding factors can have a huge impact on the results of case studies, so it is essential that the researcher defines a means to control and reduce the effects of confounding factors (it might not always be possible to eliminate them). Kitchenham et al. [19] use the example where different software engineers with different skill levels are used for testing two tools. This type of confounding factor can affect the internal validity of the study. In general, Kitchenham et al. [19] give several other examples of frequent confounding factors in software case studies:

- Learning how to use a method or tool while its benefits are being assessed. To avoid this effect, the researcher must ensure that the activities aimed at learning how to use a new method or technology are separate from those aimed at evaluating it.
- Using staff who are either very enthusiastic or very sceptical about the method or tool under evaluation. To minimize this

effect, a case study project should be undertaken using a normal staff-allocation method.
- Comparing different application types. Case studies should not compare across application domains. Appropriate selection of case-study projects should help avoid this problem.

Step 4. Ensure strategy for data validity

The four design tests for case study research are: construct validity, internal validity, external validity and reliability. At this stage, it is possible to describe an initial construct validity whereby the "constructs" are developed by the researcher using multiple sources of evidence, such as literature and expert opinion. It is not until later on that the actual constructs are tested (also using multiple sources of evidence, such as interviews, observation, documentation, etc.).

The four validity tests have been described at length in a number of books (see [45], pp.33-39); we will only mention them briefly here.

- Construct validity: establishing correct operational measures for the concepts being studied.
- Internal validity: establishing a causal relationship, whereby certain conditions are shown to lead to other conditions, as distinguished from spurious relationships.
- External validity: establishing the domain to which the study's findings can be generalized.
- Reliability: demonstrating that the operations of the study (e.g., the data collection procedures) can be replicated with the same results.

Step 5. Define the data collection strategy and process

A data collection strategy establishes what kind of evidence will be collected and how. There are six main sources of evidence: documentation, archival records, interviews, direct observations, participant observation and physical artifacts. Each type of evidence has different strengths and weakness and is applicable in different contexts [45]. The researcher must determine what type of evidence can be collected and ensure that multiple sources are used. In general, it is important that a researcher collects case study data in as many ways as possible to help strengthen the research findings. The researcher must define at this stage how the chain of evidence will be preserved. The researcher needs to consider: what are the sources of data and what is the order in which you can get it? Some example questions for documentation are: What documentation can you get? Where will you get if from? And for interviews: Who are you going to talk to and in what order? Interview questions should be formulated based on the hypotheses, concepts and measures. It is important however that there are enough open-ended questions so that the respondent can add any information he/she considers pertinent. This is particularly important for an exploratory pilot case study as open-ended questions can help the researcher identify important areas that may have been overlooked. The researcher must define at this stage how the chain of evidence will be preserved. Furthermore, the process of collecting the data must be planned for, including whether interviews will take place first, followed by observation etc.

Step 6. Design the case study plan step by step

Before commencing the case study it is important to plan and document the expected sequence of events. A logical plan should set out how the initial questions are to be answered and conclusions derived. The plan includes major phases such as the collection and analysis of relevant data and is one means to help avoid the situation where the evidence collected does not address the initial research questions. Furthermore, the researcher should also try to anticipate where deviations from the case study plan might occur and what should be done to address for such changes.

Step 6.1. Define QA on conduct of the plan

Definition of quality assurance (QA) criteria ensures that the methods or tools under investigation are used correctly and that any factors that could bias the results are recorded (such as change of staff, or a change in the priority of the case study projects). Kitchenham et al. [19] argue that it is essential that the researcher audits conformance to the experimental plan and records any changes. The purpose of the QA definition is so that at the end of the study, an evaluation report can be written up including recommendations for changes in procedures. One way to ensure that a certain level of quality is maintained throughout the case study is to seek independent feedback and review of progress at each stage. The original plan should be systematically compared with progress and results at each stage of the study.

Step 6.2. Design a case study storage system

A number of tools may be used for storing different types of case study data (theoretical frameworks, data tables and interview transcripts). A simple database system may be suitable, but more sophisticated tools such as NVivo [28] act not only as a repository for case study data, but also help the researcher organize the data, record insights and query evidence about relations between items, processes and people. Such tools are very powerful for making sense of complex qualitative data collected in case studies. In particular, NVivo provides functionality for classifying, sorting and arranging information, making it possible to explore trends, find meaning and arrive at answers to research questions.

Step 6.4. Produce the first draft of the plan

The draft of the plan should contain precise and unambiguous descriptions of, and formal documentation on: a) the chosen protocol; b) the chosen data collection strategy and process; c) the designed case study storage system; d) QA procedures on conduct of the plan; e) other steps in the case study (to the extent known at this time).

Step 7. Have the draft plan externally checked

When undertaking multiple case studies it is advisable that the researcher has the plan externally checked. This is especially beneficial to ensure that: time-estimates are realistic, ethical issues have been addressed and adequate resources are available and included in the plan. Moreover, the external reviewer should check that sufficient evidence to address the research questions will be established given the described data collection methods and the planned data analysis methods.

Step 8. Update the plan based on feedback

Any changes or improvements to the plan suggested by the

external reviewer(s) in the previous step should be reflected in an updated version of the plan.

Step 9. Undertake pilot case study

This will allow the researcher to check that the research propositions still make sense and that the interview questions are actually addressing the propositions. If there are any problems with the interview questions they may need to be modified. The pilot study may identify if an already developed framework needs modification. Pilot studies in software engineering were discussed by Glass in [14].

VI. Data Collection

The boundary between data collection, data processing, data analysis, data interpretation, data presentation, and reporting can at times become somewhat blurred. Data triangulation, which must never be lost sight of during the data collection stage, crosscuts data processing and data interpretation. The three principles of data collection are: 1) use multiple sources of evidence, 2) create a case study database, and 3) validate data and maintain a chain of evidence [22, 23, 26, 44, 45]. These three principles correspond to the following three steps in the data collection phase.

Step 1: Obtain the data from multiple sources

Evidence for case studies may come from six main sources: documents, archival records, interviews, direct observation, participant-observation and physical artifacts [36, 45]. However, other sources are possible, such as films, photography, ethnography and life histories. We now describe the six most common forms of case study evidence:

a) Interviews

Interviews are essential sources for the collection of case study data. The advantages of using interviews as a form of data collection are that they focus directly on the case study topic and can provide perceived casual inferences. Interview preparation includes carefully constructing a set of interview questions, determining the sampling strategy and determining the number of interviewees and interviewers. Outputs of the interviewing process are: interview transcriptions, interview notes and any scheduled follow-up meetings. The researcher must allow for deviations to the interview schedule. For example, during the interviews, respondents may recommend other key persons who should be interviewed and may recommend other relevant sources of evidence.

Yin [45] recommends that the case study researcher should i) follow the line of inquiry reflected in the case study protocol, and ii) during the interview ask questions in an unbiased manner, that at the same time, serve the specific line of inquiry. Interviews may be constructed using an open-ended or focused approach, but typically case study interviews have many questions of an open-ended nature [45]. A focused interview is usually a short interview (approximately one hour) where a certain set of questions are followed from the case study protocol. The major purpose of such interviews is to corroborate certain facts [44, 45].

Key weaknesses that may arise from using interviews are: bias due to poorly constructed questions, response bias, inaccuracies due to poor recall and reflexivity – where the respondent says what the interviewer expects or wants to hear [45]. These weaknesses must be addressed by: 1) carefully constructing the questions and having them reviewed by an external reviewer; 2) interviewing an appropriate number of interviewees and ensuring other sources of evidence are used; 3) allowing the interviewee time to express her or his opinion about certain events or procedures and to describe the actual facts. Writing up the interview transcripts should be started as soon as possible after the interviews. As mentioned previously, software such as NVivo [28] can help tremendously with documenting interview transcript data. The final transcription report is stored and held as part of the case study database. It is very difficult to take notes, keep one's place in the list of questions (particularly when an interviewee may answer several unasked questions at one time), and preserve a professional demeanour. Therefore, we recommend recording each interview, if the interviewee agrees to this.

b) Archival records

The case study researcher may also undertake a search of an organization's archival data as another useful source of evidence. Examples of archival records are: computer files, logs and records, organizational charts; service, personnel or financial records. Archival records are useful because they tend to be precise and can provide some quantitative evidence for analysis. Thus, they complement and should be used in conjunction with other sources of information in producing a case study. However, often archival data is confidential and accessibility is limited. The case study researcher must be careful when interpreting the usefulness and accuracy of records.

c) Documentation

Documentary information is another important source of evidence in a case study. Supplementary documents play an explicit role in data collection throughout the study. Examples of such documentation are: requirements documents, system documentation, user manuals, contracts, letters, memoranda, meeting minutes, proposals, progress reports, formal studies, evaluations and newspaper clippings. As documentation becomes quickly out of date, the case study researcher must be careful not to rely too heavily on such documents. In addition, it may be difficult to determine the reporting bias of the author who compiled the document. A documentation report should be produced by the case study researcher describing and classifying the different types of documents to be analysed and how they may link to other sources of evidence.

d) Direct observation

Another useful data collection method is direct observation, i.e., observing events, tasks, etc., to find facts and other valuable information. Direct observation occurs when a researcher visits the case study "site" and observes certain events or phenomena relevant to the line of inquiry. The researcher will take notes and may record "actual" events. Text or numerical data may result and the events and their context should be recorded. The researcher will be able to obtain detailed and accurate information about the people she/he is studying. Ob-

servable details (like daily time allotment) are more easily observed and understandable over a longer period of time. A strength of using observation is that researchers can discover discrepancies between what participants say - and often believe - should happen (the formal system) and what actually does happen, or between different aspects of the formal system. In contrast, a one-time survey of people's answers to a set of questions might be quite consistent, but is less likely to show conflicts between different aspects of the social system or between conscious representations and behaviour [39]. Possible observation threats for participant observation (discussed in the next paragraph) are also relevant for direct observation (albeit to a somewhat lesser degree).

e) Participant observation

The aim of participant observation is to gain a close and intimate familiarity with a given group of individuals or a particular community and their practices through *involvement* with people in their natural environment, often, though not always, over an extended period of time [39]. Although this method is generally characterized as qualitative research, it can include quantitative dimensions. Increased threats to validity and a downside of participant observation are: 1) the objectivity of the researcher, and 2) possible observer effects, i.e, observation and participation may distort the observed behaviour.

f) Physical artifacts

There may be physical artifacts (such as computer software and hardware, office layout and equipment etc.) that may be useful as a data source, particularly for validation of certain facts and views gleaned in interviews or observations. For example, poor working conditions may be a contextual factor that may have an impact on project outcomes.

Step 2: Store the data into the database

Input to this step is the raw material (including interview transcripts, the researcher's field notes, documents collected during data collection, and survey material) and the output is the case study database which will hold all facts and evidence. The data may consist of coded data; a coding scheme; memos and other analytic material; and data displays.

Step 3: Validate the data

The data collected should be validated and a data validation report prepared. This should discuss the actual data collection and report any deviations from the data collection plan. The researcher should check if any data is missing or invalid and must ensure that the collected data is sufficient to answer questions and propositions of the case study.

If data is: a) missing; b) wrong; c) invalid; d) or not enough, or if multiple sources of evidence cannot support data triangulation, the researcher will need to return to the collection of data. If quantitative and qualitative evidence do addresses the questions and propositions of the case study, then the data should be stored into the database.

VII. Data Analysis

In this phase we must ensure that a complete chain of evidence is kept, alternate perspectives and explanations are considered, and that we develop clear conclusions for practice/further research. We found that a useful method for maintaining a chain of evidence is to highlight and number important material in interview transcriptions. This material can then be numbered, copied, inserted into tables, and sorted. The sorting then helps with cross-case analysis. The numbering ensures that the material can be traced directly back to its source; i.e., the actual interview and position within the interview. The transcriptions can then be included in an appendix. At this stage must consider if there are threats to validity and address them in a systematic way.

When multiple case studies are involved data must be analysed both within case and across case [11, 45]. This analysis usually consists firstly of a detailed analysis of each case separately, followed by a comprehensive cross case analysis. If there is sufficient data, appropriate statistical techniques should be used. If the data is continuous then parametric statistics will be appropriate (e.g., ANOVA, linear regression). However, if the data is categorical then non-parametric statistical techniques must be used (e.g., chi square, logistic regression). Non-parametric methods are widely used for studying populations that take on a ranked order (such as software reviews receiving one to four stars) and may be necessary when data has a ranking but no clear numerical interpretation, such as when assessing preferences [39].

VIII. Reporting

We must ensure that any reporting is easy to read, is suitable for its audience, and is written in an engaging manner. In addition to reporting the research questions and answers, case study cause and context should be adequately covered. Further, related theory, hypotheses and propositions must be clear, and data collection procedures should be presented with inclusion of adequate raw data. Sufficient evidence should be displayed, and it must be made clear how the evidence was selected for inclusion. Additionally, conclusions and implications for practice and future research should be covered [36, 40, 45].

It is crucial to identify the audience, as the report should be tailored to its prospective audience. This may include: academic colleagues, policy makers, practitioners, thesis committees or examiners and funders of research. With many different audiences, the researcher may find that several different reports may be required. Various good practices, such as writing up while the case study is conducted (which reduces effort in the long run) can help the reporting process. We advise to start writing as soon as possible – the literature and design can be reported in the protocol. The methodology section can also be started during the case study focusing or planning phase. Descriptive data can be reported prior to analysis. The case study report must be complete with the boundaries between case and context defined; all relevant evidence must have been collected; the case study should have been completed to a predefined schedule and not finished purely because of time constraints.

The structure of the report has to fit with the case study design and the needs of the reader. There are three basic formats:

1) single case study narrative report; 2) a multiple case study comprising a narrative report for each case; and, 3) a cross-case analysis with a series of questions and answers based on the relationship between questions and evidence. For multiple case studies, naturally, formats 2) and 3) are appropriate. The researcher should report in a well-defined sequence and may choose a sequence such as: 1) issue or problem, 2) review of prior literature, 3) research methods used, 4) findings from data collected and analysed, 5) conclusions and implications. This type of sequence is normally used for journal articles. The researcher may need to use some kind of comparative structure, as she/he has repeated the case study two or more times, and should compare alternative descriptions and their explanations. The structure may need to be chronological as longitudinal case studies are normally presented in chronological order. The researcher may need to consider explanations when investigating cause-effect propositions; or if she/he is conducting theory building each section may reveal a new part of the theoretical argument being made; or she/he may choose an unsequenced structure where there is no particular order; however, this latter organization is usually best reserved for descriptive case studies. The researcher must consider all alternative perspectives, i.e. rival propositions must be discussed in the report.

It is necessary to ask the individuals who provided the information to confirm that all the information is correct. Anonymity of case study participants will depend on agreements with the organization and the individuals involved (from the case study administration phase). If possible, anonymous cases and individuals should be avoided. If cases and individuals are anonymous, background context will be lost. An independent review can be helpful for identifying logical flaws or identifying alternative explanations that have been omitted.

IX. Further Work

We have tested our case study guidelines in three different exploratory research projects in which we have been involved. The guidelines have been modified based on the feedback we have received. Feedback resulted mainly in clarifications and modifications to the ordering of some of the activities. Our case study guidelines were found to be very useful for exploratory case studies in software engineering. Some of our students commented that having the guidelines has made the writing of their methodology chapter much easier. We have adopted the guidelines within our research group as a standard for use with exploratory research.

Our guidelines have not yet been tested in explanatory or evaluatory case studies. However, this is something that we hope to do in the future. We would be pleased if other researchers used our guidelines and provided us with feedback so that we can modify the guidelines as appropriate. Our aim is to ensure the guidelines are helpful for software engineering researchers and provide them with a useful framework when the case study research methodology is appropriate.

ACKNOWLEDGEMENTS

The authors would like to thank the students who have been testing the guidelines in their research, as well as the other researchers who have provided us with much useful feedback.

REFERENCES

[1] Atkins, C., & Sampson, J., "Critical appraisal: guidelines for single case study research", European Conference on information Systems, June 6-8, Gdańsk, Poland, 2002, pp100-109.

[2] Benbasat, I., D. K Goldstein, and M. Mead. "The case research strategy in studies of information systems." *MIS Quarterly* 11 (1987): 369-86.

[3] Bouwman H., Faber E, 2003; "Case study protocol B4U" Technical Report, Telematica Instituut, TU Delft, TI/RS/2003/008, Downloaded 27th October, 2008 https://doc.freeband.nl/dsweb/Get/Document-29482/D3.1%20case%20study%20protocol%20B4U.pdf.

[4] Brereton P Kitchenham Bm and Budgen D "Using a protocol template for case study planning" EASE, 2007

[5] Cavaye, A.L.M., "Case study research: A multi-faceted research approach for IS", Information Systems Journal, Vol. 6, 1996, pp227-242.

[6] Creswell, J.W., Qualitative inquiry and research design: Choosing among five traditions, Sage, Thousand Oaks, CA, 1998.

[7] Darke, P., Shanks, G., & Broadbent, M., "Successfully completing case study research: Combining rigor, relevance and pragmatism", Information Systems Journal, Vol. 8, No. 4, 1998, pp273-289.

[8] de Weerd-Nederhof, P.C., "Qualitative case study research: The case of a PhD research Project on organizing and managing new product development systems", Management Decision, Vol. 39, No. 7, 2001, pp513-538.

[9] Dubé, L., & Paré, G., "Rigor in information systems positivist case research: Current practices, trends, and recommendations", MIS Quarterly, Vol. 27, No. 4, December 2003, pp597-635.

[10] Dubin R, Theory building, Free Press, New York 1978

[11] Eisenhardt, K.M., "Building theories from case study research", Academy of Management Review, Vol. 14, No. 4, October 1989, pp532-550 (see also rejoinder by Dyer & Wilkins, 1991, and response by Eisenhardt, 1991).

[12] Eisenhardt, K.M., "Better stories and better constructs: The case for rigor and comparative logic", Academy of Management Review, Vol. 16, No. 3, July 1991, pp629-627 (response to Dyer & Wilkins, 1991).

[13] Flyvbjerg Bent "Five misunderstandings about case-study research", Qualitative Inquiry, Vol. 12, No. 2, 219-245 (2006).

[14] Glass, R. L., "Pilot studies: What, why, and how", Journal of Systems Software, Vol. 36, No. 1, January 1997, pp85-97.

[15] Host M and Runeson P "Case studies on software engineering programmes" Retrieved 27/10/2008 http://serg.telecom.lth.se/research/publications/docs/420_Technical%20report.pdf.

[16] Jauch, L.R., Osborn, R.N., & Martin, T.N., "Structured content analysis of cases: A complementary method of organizational research", Academy of Management Review, Vol. 5, No. 4, October 1980, pp517-525.

[17] Jensen, J.L., & Rodgers, R.,"Cumulating the intellectual gold of case study research", Public Administration Review, Vol. 61, No. 2, March/April 2001, pp235-246.

[18] Kaplan, B., & Duchon, D., "Combining qualitative and quantitative methods in information systems research: A case study", MIS Quarterly, Vol. 12, No. 4, December 1988, pp571-586.

[19] Kitchenham, B.A., Pickard, L.M., & Pfleeger, S.L., "Case studies for method and tool evaluation", IEEE Software, Vol. 12, No. 4, July 1995, pp52-62.

[20] Kitchenham, B.A., & Pickard, L.M., "Evaluating software engi-

neering methods and tools – Part 9: Quantitative case study methodology", Software Engineering Notes, Vol. 23, No. 1, January 1998, pp24-26.

[21] Kitchenham, B.A., & Pickard, L.M., Evaluating software engineering methods and tools – Part 10: Designing and running a quantitative case study", Software Engineering Notes, Vol. 23, No. 3, May 1998, pp20-22.

[22] Klein, H.K., & Myers, M.D., "A set of principles for conducting and evaluating interpretive field studies in information systems, MIS Quarterly, Vol. 23, No. 1, March 1999, pp67-93.

[23] Lee, A.S., "A scientific methodology for MIS case studies", MIS Quarterly, Vol. 13, No. 1, March 1989, pp33-50.

[24] Lee, A.S., "Case studies as natural experiments", Human Relations, Vol. 42, No. 2, 1989, pp117-137.

[25] Leonard-Barton, D., "A dual methodology for case studies: Synergistic use of a longitudinal single site with replicated multiple sites", Organization Science, Vol. 1, No. 3, August 1990, pp248-266.

[26] Lethbridge, T.C., Sim, S.E., & Singer, J., "Studying software engineers: Data collection techniques for software field studies", Empirical Software Engineering, Vol. 10, No. 3, 2005, pp311-341.

[27] Lincoln, Y.S., & Guba, E.G., "Judging the quality of case study reports", in Huberman, A.M., & Miles, M.B., (Eds.), The Qualitative Researcher's Companion, Sage Publications, Thousand Oaks, CA, 2002, pp205-215 (reprinted from Qualitative Studies in Education, Vol. 3, No. 1, 1990, pp53-59).

[28] NVivo QSR web site www.qsrinternational.com

[29] Paré, G., & Elam, J.J., "Using case study research to build theories of IT implementation", in Lee, A.S., Liebenau, J., & DeGross, J.I., (Eds.), Information Systems and Qualitative Research: Proceedings of the IFIP TC8 WG 8.2 International Conference on Information Systems and Qualitative Research, 31st May – 3rd June 1997, Philadelphia, Pennsylvania, USA, Chapman & Hall, London, 1997, pp542-568.

[30] Paré, G., "Investigating information systems with positivist case study research", Communications of the AIS, Vol. 13, 2004, pp233-264.

[31] Perry, D.E., Sim, S.E., & Easterbrook, S.M., "Case studies for software engineering", Proceedings of the 26th International Conference on Software Engineering (ICSE'04), 2004.

[32] Platt, J., "What can case studies do?", Studies in Qualitative Methodology, Vol. 1, 1988, pp1-23.

[33] Scheepers, R., & Scheepers, H., "Contexts of relevance in exploratory case studies in information systems: Ubiquitous information technology implementation in organizations", in S.T. March, A.P. Massey, & DeGross, J.I., (Eds.), Proceedings of the Twenty-Fourth International Conference on Information Systems, December 14-17, 2003, Seattle, WA, 2003, pp25-35.

[34]Shanks, G., & Parr, A., "Positivist, single case study research in information systems: A critical analysis", Proceedings of the European Conference on Information Systems, Naples, June 2003.

[35] Smith, N.C., "The case study: A useful research method for information management", Journal of Information Technology, Vol. 5, No. 3, 1990, pp123-133.

[36] Stake, R.E., The art of case study research, Sage, Thousand Oaks, CA, 1995.

[37] Walsham, G., "Interpretive case studies in IS research: Nature and method", European Journal of Information Systems, Vol. 4, 1995, pp74-81.

[38] White H, "Introduction to BPMN", BP Trends July 2004, p1-11 www.bptrends.com

[39] Wilkipedia http://en.wikipedia.org/wiki/Participant_observation retrieved Oct 24, 2008.

[40] Woodside, A.G., & Wilson, E.J., "Case study research methods for theory building", Journal of Business & Industrial Marketing, Vol. 18, No. 6-7, 2003, pp493-508.

[41]Yin, R.K., "Research design issues in using the case study method to study management information systems", Harvard Business School Research Colloquium, 1989, pp1-5 (in Cash Jr., J.I., & Lawrence, P.R., (Eds.), The Information Systems Research Challenge: Qualitative Research Methods, Vol. 1, Harvard Business School, Boston, MA, 1991, pp1-6).

[42]Yin, R.K., "The case study crisis: Some answers", Administrative Science Quarterly, Vol. 26, No. 1, March 1981, pp58-65.

[43] Yin, R.K., "Enhancing the quality of case studies in health services research", Health Services Research, Vol. 34, No. 5, December 1999, p1209-1224.

[44] Yin, R.K., Applications of case study research, Second Edition, Sage Publications, Thousand Oaks, CA, (1st Edition, 1993), 2003.

[45] Yin, R.K., Case study research: Design and methods, Third Edition, Sage Publications, Thousand Oaks, CA, (2nd Edition, 1994), 2003.